Least Squares Symmetrizable Solutions for a Class of Matrix Equations

Fanliang Li
School of Sciences, Institute of Mathematics and Physics, Central South University of Forestry and Technology, Changsha, China
Email: lfl302@tom.com

Received January 30, 2013; revised March 27, 2013; accepted April 4, 2013

Copyright © 2013 Fanliang Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

In this paper, we discuss least squares symmetrizable solutions of matrix equations \((AX = B, XC = D)\) and its optimal approximation solution. With the matrix row stacking, Kronecker product and special relations between two linear subspaces are topological isomorphism, and we derive the general solutions of least squares problem. With the invariance of the Frobenius norm under orthogonal transformations, we obtain the unique solution of optimal approximation problem. In addition, we present an algorithm and numerical experiment to obtain the optimal approximation solution.

Keywords: Matrix Equations; Matrix Row Stacking; Topological Isomorphism; Least Squares Solution; Optimal Approximation

1. Introduction

The matrix equations \((AX = B, XC = D)\), where \(A, B, C, D\) are usually given by experiments, have a long history [1]. Many authors considered these matrix equations. For example, Mitra [2,3], Chu [4] discussed its unconstraint solutions with the generalized inverse of matrix and the singular value decomposition (SVD), respectively. In recent years, many authors considered its constraint solutions. A series of meaningful results were achieved [1,5-10]. The methods in these papers are mainly the generalized inverse of matrix and special properties of finite dimensional vector spaces [1,5,6], the decomposition of matrix or matrix pairs [7,8] and the special properties of constraint matrices [9,10]. However, the least squares symmetrizable solutions for these matrix equations have not been considered. The purpose of this paper is to discuss its least squares symmetrizable solutions with the matrix row stacking, Kronecker product and special relations between two linear subspaces which are topological isomorphism because the structure of symmetrizable matrices can not be found and the methods applied in [1-10] can not solve the problem in this paper. The background for introducing the definition of symmetrizable matrices is to get “symmetric” matrices from nonsymmetric matrices [11] because of nice properties and multi-areas of applications of symmetric matrices. For example, Sun [12] introduced the definition of positive definite symmetrizable matrices to study an efficient algorithm for solving the nonsymmetry second-order elliptic discrete systems.

Throughout this paper we use some notations as follows. Let \(R^{n \times m}\) be the set of all \(n \times m\) real matrices and denote \(R^n = R^{n \times 1}\); \(OR^{n \times m}\), \(SR^{n \times m}\) and \(ASR^{n \times m}\) are the set of all \(n \times n\) orthogonal, symmetric and skew-symmetric matrices, respectively. \(R(A), A^T\) and \(A^\dagger\) represent the range, the transpose and the Moore-Penrose generalized inverse of \(A\), respectively.

\(I_n\) denotes the identity matrix of order \(n\). For \(A = \begin{pmatrix} a_{ij} \end{pmatrix}\), \(B = \begin{pmatrix} b_{ij} \end{pmatrix} \in R^{m \times n}\), \(A \otimes B\) denotes Kronecker product of matrix \(A\) and \(B\); \(\langle A, B \rangle = tr(B^TA)\) denotes the inner product of matrix \(A\) and \(B\). The induced matrix norm is called Frobenius norm, i.e. \(\|A\| = \langle A, A \rangle^{1/2}\), then \(R^{n \times m}\) is a Hilbert inner product space.

Definition 1. A real \(n \times n\) matrix \(A\) is called a symmetrizable (skew-symmetrizable) matrix if \(A\) is similar to a symmetric (skew-symmetric) matrix \(A\). The set of symmetrizable (skew-symmetrizable) matrices is denoted by \(SZR^{n \times n}\) (\(ASZR^{n \times n}\)).

From Definition 1, it is easy to prove that \(A \in SZR^{n \times n}\) (\(ASZR^{n \times n}\)) if and only if there exists a nonsingular matrix \(W\) and a symmetric (skew-symmetric) matrix \(\tilde{A}\) such that \(A = W\tilde{A}W^{-1}\).

We now introduce the following two special classes of
subspaces in $R^{n\times n}$.

$$SR^{n\times n}(W) = \left\{ A \mid A = W\overline{A}W^{-1}, \forall A \in SR^{n\times n} \right\},$$

$$ASR^{n\times n}(W) = \left\{ A \mid A = W\overline{A}W^{-1}, \forall A \in ASR^{n\times n} \right\}.$$  

It is easy to see that if $W$ is a given nonsingular matrix, then $SR^{n\times n}(W)$ and $ASR^{n\times n}(W)$ are two closed linear subspaces of $R^{n\times n}$. In this paper, we suppose that $W$ is a given nonsingular matrix and $W \neq I_n$. We will consider the following problems.

**Problem I.** Given $A, B \in R^{k\times n}$, $C, D \in R^{n\times l}$, find $X \in SR^{n\times n}(W)$ such that

$$\min \left\{ \|AX - B\|^2 + \|XC - D\|^2 \right\}.$$  

**Problem II.** Given $X^* \in R^{n\times n}$, find $\hat{X} \in S_\epsilon$ such that

$$\|X^* - \hat{X}\| = \min_{X \in S_\epsilon} \|X^* - X\|,$$

where $S_\epsilon$ is the solution set of Problem I.

In this paper, if $C = 0, D = 0$ in Problem I, then Problem I becomes Problem I of [13]. Peng [13] studied the least squares symmetrizable solutions of the matrix equation $AX = B$ with the singular value decomposition of matrix. The method applied in [13] can not solve Problem I in this paper. In this paper, we first take matrix equations $(AX = B, XC = D)$ into linear equations with matrix row stacking and Kronecker product. Then we obtain an orthogonal basis-set for $SR^{n\times n}(W)$ with special relations between two linear subspaces which are topological isomorphism. Based on these results, we obtain the general expression of Problem I.

This paper is organized as follows. In Section 2, we first discuss the matrix row stacking methods, Kronecker product of matrix and relations between $SR^{n\times n}(W)$ and $SR^{n\times n}$. Then we obtain the general solutions of Problem I. In Section 3, we derive the solution of Problem II with the invariance of Frobenius norm under orthogonal transformations. In the end, we give an algorithm and numerical experiment to obtain the optimal approximation solution.

### 2. The Solution Set of Problem I

At first, we discuss the matrix row stacking methods, Kronecker product of matrix and relations between two linear subspaces which are topological isomorphism.

For any $A \in R^{n\times n}$, let $vec(A)$ denote an ordered stack of the row of $A$ from upper to low stacking with the first row, i.e.

$$vec(A) = \left[ A(1,:)A(2,:),\cdots,A(m,:) \right]^T \in R^{mn},$$  

where $A(i,:)$ denotes the $i$th row of $A$. For any vector $x \in R^{mn}$, let $vec^{-1}(x)$ denote the following matrix containing all the entries of vector $x$.

where $x(i:j)$ denotes the elements from $i$ to $j$ of vector $x$. From (2.1), we can derive the following two linear equations between two linear subspaces of $R^{n\times n}$.

$$\vec{S}(SR^{n\times n}(W)) = \left\{ \vec{vec}(A) \mid A \in SR^{n\times n} \right\},$$

$$\vec{S}(SR^{n\times n}(W)) = \left\{ \vec{vec}(A) \mid A \in ASR^{n\times n}(W) \right\}.$$  

**Lemma 1.** [14] If $A \in R^{k\times n}, B \in R^{n\times l}, X \in R^{n\times n}$, then

$$\vec{S}(AXB) = (A \otimes B^T)\vec{S}(X).$$  

For any $X \in SR^{n\times n}$, let

$$f^*: X \in SR^{n\times n}, \text{i.e.} f^*(X) = W\overline{X}W^{-1}.$$  

It is easy to prove that mapping $f^*$ is a topological isomorphism mapping from $SR^{n\times n}$ to $SR^{n\times n}(W)$. According to (2.1) and (2.5), it is easy to derive the following mapping from linear subspaces $vec(SR^{n\times n})$ to $vec(SR^{n\times n}(W))$.

$$f : vec(X) \mapsto (W \otimes W^{-T})vec(X),$$

i.e.

$$f(vec(X)) = (W \otimes W^{-T})vec(X).$$  

It is also easy to prove that mapping $f$ is a topological isomorphism mapping from $vec(SR^{n\times n})$ to $vec(SR^{n\times n}(W))$. It is clear that the dimension of $SR^{n\times n}$ is $\frac{n(n+1)}{2}$. This implies that the dimension of $vec(SR^{n\times n})$ and $vec(SR^{n\times n}(W))$ are also $\frac{n(n+1)}{2}$. In this paper, let $k = \frac{n(n+1)}{2}$.

**Lemma 2.** If $\alpha_1, \alpha_2, \cdots, \alpha_k$ is an orthonormal basis-set for $vec(SR^{n\times n}(W))$, and let $\Lambda = (\alpha_1, \alpha_2, \cdots, \alpha_k)$, then the following relations hold.

$$\Lambda \in R^{\frac{n(n+1)}{2}}, \Lambda^T\Lambda = I_k,$$

$$R(\Lambda) = vec(SR^{n\times n}(W)), R(\Lambda^T) = R^k.$$  

From the definition of the orthonormal basis-set, it is easy to prove Lemma 2, so the proof is omitted.

For any matrix $A \in SR^{n\times n}(W)$, if let $vec(A) \in R^k$, denote the vector of coordinates of $vec(A)$ with respect to the basis-set $\alpha_1, \alpha_2, \cdots, \alpha_k$, then combining (2.4) and
(2.7), we have

\[ \text{vec}(A) = \Lambda \text{vec}(A) + \epsilon, \] (2.8)

where \( \Lambda \) is a diagonal matrix and \( \epsilon \) is a vector.

Moreover, for any \( x \in R^k \), the following conclusion holds.

\[ \text{vec}^{-1}(Ax) \in SR^{\alpha,\epsilon}(W). \] (2.9)

This implies that finding \( X \in SR^{\alpha,\epsilon}(W) \) such that \( \min \{\|AX - B\| + \|XC - D\|\} \) if and only if finding \( \alpha \in R^k \) such that

\[ \min \left\{ \left\| A \otimes I_n \Lambda \alpha - \left( \frac{\text{vec}(B)}{\text{vec}(D)} \right) \right\|^2, \right\} \] (2.13)

where \( \Lambda \alpha = \text{vec}(X) \). From Lemma 3, the general solutions of (2.13) is

\[ \alpha = A_0^\ast B_0 + \left( I_k - A_0^\ast A_0 \right) K, \forall K \in R^k. \] (2.14)

Combining (2.13) and (2.14) gives (2.12). \( \square \)

### 3. The Solution of Problem II

Let \( S_E \) be the solution set of Problem I. From (2.12), it is easy to see that \( S_E \) is a nonempty closed convex set. So we claim that for any given \( X^* \in R^{\alpha,\epsilon} \), there exists the unique optimal approximation for Problem II.

**Theorem 2.** If given \( A, B \in R^{\alpha,\epsilon} \), \( C, D \in R^{k,\epsilon} \), then Problem II has a unique solution \( \hat{X} \in S_E \).

Moreover, \( \hat{X} \) can be expressed as

\[ \hat{X} = \text{vec}^{-1}(A \hat{\alpha}), \]

\[ \hat{\alpha} = A_0^\ast B_0 + \left( I_k - A_0^\ast A_0 \right) \Lambda^T \text{vec}(X), \] (3.1)

where \( A_0, B_0 \) are denoted by (2.11).

**Proof.** Choose \( \hat{\Lambda} \) such that \( (\Lambda, \hat{\Lambda}) \in OR^{\alpha,\epsilon} \). Combining the invariance of the Frobenius norm under orthogonal transformations, (2.12) and (2.14), we have

\[ \|X^* - \hat{X}\|^2 = \|\text{vec}(X^*) - \text{vec}(\hat{X})\|^2 = \|\text{vec}(X^*) - \Lambda \hat{\alpha}\|^2 = \left( \Lambda^T \text{vec}(X^*) - \hat{\Lambda} \right) \Lambda^T \hat{\alpha} \]

\[ = \|\Lambda^T \text{vec}(X^*) - A_0^\ast B_0 - (I_n - A_0^\ast A_0) K\|^2 + \|\hat{\Lambda} \text{vec}(\hat{X})\|^2. \]

Let \( E = I_k - A_0^\ast A_0 \), \( E_i = I_k - E \), it is clear that \( E, E_i \) are orthogonal projection matrices satisfying \( EE_i = 0 \). Hence, we have

\[ \|X^* - \hat{X}\|^2 = \|E \left( \Lambda^T \text{vec}(X^*) - A_0^\ast B_0 - K \right)\|^2 + \|E_i \left( \Lambda^T \text{vec}(X^*) - A_0^\ast B_0 \right)\|^2 + \|\hat{\Lambda} \text{vec}(\hat{X})\|^2. \]

It is easy to prove that \( E \Lambda^T \text{vec}(X^*) - A_0^\ast B_0 = 0 \). This implies that

\[ \min_{\forall x \in S_E} \|X^* - \hat{X}\| \Leftarrow \min_{\forall x \in S_E} \|E \left( \Lambda^T \text{vec}(X^*) - A_0^\ast B_0 - K \right)\|. \] (3.2)
The solution of (3.2) is
\[ K = \Lambda^T \text{vec}(X^*) + E_i K_i, \quad \forall K_i \in R^i. \]  
(3.3)
Substituting (3.3) to (2.12) gives (3.1). □
From Theorem 2, we can design the following algorithm to obtain the optimal approximate solution.

**Algorithm**
1) Input \( A, B, C, D, W, X^* \).
2) Input a basis-set \( A_i, A_{i+1}, \ldots, A_s \) for \( SR^{n\times n} \).
3) According to the calculation procedure before, compute \( \alpha_1, \alpha_2, \ldots, \alpha_s \) and obtain an orthonormal basis-set for \( \text{vec}(SR^{n\times n}(W)) \).
4) Let \( L = A_1, A_2, \ldots, A_s \), compute \( A_0, B_0 \) from (2.11).
5) Compute \( \hat{\alpha} \) from the second equation of (3.1).
6) According to the first equation of (3.1), calculate \( \hat{X} \).

**Example** \((n=3, h=4, l=5)\)
1) Input \( A, B, C, D, W, X^* \) as follows.
\[
A = \begin{bmatrix}
1.9 & 5.4 & 6.5 \\
2.8 & 7.8 & 3.9 \\
3.6 & -0.9 & 7.5 \\
-2.7 & 2.1 & -3.9
\end{bmatrix},
\]
\[
B = \begin{bmatrix}
5.9 & 12.3 & -7.5 \\
10.3 & 4.1 & 2.9 \\
-3.7 & 5.6 & 3.7 \\
5.1 & 1.3 & -8.5
\end{bmatrix},
\]
\[
C = \begin{bmatrix}
4.1 & 5.6 & 5.4 & 3.1 & 1.6 \\
3.2 & -1.9 & 7.6 & 7.8 & 3.9 \\
1.9 & 2.8 & 4.3 & -2.3 & 2.7
\end{bmatrix},
\]
\[
D = \begin{bmatrix}
-1.3 & 7.6 & 6.7 & -2.1 & 2.9 \\
5.9 & 2.3 & 1.8 & 3.1 & 5.6 \\
4.5 & 5.4 & 6.7 & 5.7 & 7.3
\end{bmatrix},
\]
\[
W = \begin{bmatrix}
1 & 2 & 3 \\
2 & 2 & 1 \\
3 & 4 & 3
\end{bmatrix},
\]
\[
X^* = \begin{bmatrix}
5.8 & 1.7 & 2.9 \\
-5.9 & -2.5 & 8.5 \\
-2.4 & 3.3 & 7.6
\end{bmatrix}.
\]
2) Input a basis-set \( A_1, A_2, \ldots, A_s \) for \( SR^{3\times3} \) as follows.
\[
A_1 = \begin{bmatrix}
0 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0
\end{bmatrix},
A_2 = \begin{bmatrix}
0 & 1 & 0 \\
1 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix},
A_3 = \begin{bmatrix}
0 & 0 & 1 \\
0 & 0 & 0 \\
1 & 0 & 0
\end{bmatrix},
A_4 = \begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix},
\]
3) According to the calculation procedure, we obtain an orthonormal basis-set \( \alpha_1, \alpha_2, \ldots, \alpha_s \) for \( \text{vec}(SR^{3\times3}(W)) \) as follows.
\[
\alpha_1 = \begin{bmatrix}
-0.0714 \\
-0.2143 \\
0.1429 \\
-0.1429
\end{bmatrix},
\alpha_2 = \begin{bmatrix}
0.2618 \\
0.4286 \\
-0.2143 \\
0.6618
\end{bmatrix},
\alpha_3 = \begin{bmatrix}
0.4286 \\
-0.1396 \\
-0.0642 \\
-0.1215
\end{bmatrix},
\alpha_4 = \begin{bmatrix}
-0.0910 \\
-0.0976 \\
-0.3984 \\
-0.3584
\end{bmatrix},
\alpha_5 = \begin{bmatrix}
0.2618 \\
0.4286 \\
-0.2143 \\
0.6618
\end{bmatrix},
\alpha_6 = \begin{bmatrix}
-0.2590 \\
0.2524 \\
0.2674 \\
-0.2590
\end{bmatrix}
\]
4) Using the software “MATLAB”, we obtain the unique solution \( \hat{X} \) of Problem II.
\[
\hat{X} = \begin{bmatrix}
-0.3657 & 0.1512 & 0.5837 \\
0.6314 & 0.4846 & -0.2603 \\
0.5852 & 0.6459 & 0.0550
\end{bmatrix}.
\]

**4. Conclusion**
In this paper, we first derive the least squares symmetrizable solutions of matrix equations \( AX = B, XC = D \) with the matrix row stacking and the theory of topological isomorphism, i.e. Theorem 1. Then we give the unique optimal approximation solution, i.e. Theorem 2. Based on Theorem 1 and 2, we design an algorithm to find the optimal approximation solution. Compare to [1-10], this paper has two important achievements. One is we apply the topological isomorphism theory to obtain the least squares symmetrizable solutions of matrix equations \( AX = B, XC = D \), and provide a method to solve the matrix equation, where the construct of constraint matrix can not be found. The other is we present a stable calculation procedure to obtain an orthonormal basis-set for
\( \text{vec}(SR^{\text{vec}}(W)) \), and solve the key problem of the algorithm.

5. Acknowledgements

The authors are very grateful to the referee for their valuable comments, and also thank for his helpful suggestions.

This research was supported by National natural Science Foundation of China (31170532).

REFERENCES


